

청년통장 데이터 분석을 기반으로 한 청년정책 개선방안 제안

(본 스토리는 2021 데.멘.토 프로그램 참여 팀의 분석 글입니다.)

[지역경제 플랫폼]

최근 청년문제가 사회적인 이슈로 대두되고 있는 만큼 중앙정부 및 지방자치단체에서 2030 청년층을 겨냥한 여러 정책들을 제시하고 있다. 그러나 지난 2019년 '시사오늘' 기사에서 청년 대변인단을 대상으로 '청년 정책의 문제점은 무엇인가?'를 질문한 결과, 다양한 환경을 반영하지 못하여 실효성이 부족하다거나 청년의 실제 삶을 반영하지 않은 정책이라는 등 현재 시행되고 있는 청년 정책의 아쉬운 부분에 대한 지적이 있었다. 따라서 MZ세대 당사자로서 청년 정책 관련 데이터 분석을 통해 정책 수요자인 청년층의 실질적인 현황을 파악하고 개선점을 제시하여 향후 청년 정책 수립에 도움이 되고자 한다.

작성자 : MZ데멘티팀(권형준, 윤준현, 차연주, 황지영)

<목차>

- 1) 데이터 소개
- 2) EDA (탐색적 데이터 분석)
- 3) 클러스터링
- 4) 텍스트마이닝
 - ① 클러스터별 단어 네트워크 분석을 통한 청년 페르소나 제시
 - ② 클러스터별 DMR 토픽 모델링
- 5) 결론
 - ① 청년 페르소나를 기반으로 한 정책 타겟층 도출
 - ② 20대 여성 창업에 대한 정책 현황 분석 및 개선점 제안

1) 데이터 소개

- 청년 노동자 통장

도내 청년들의 꿈을 응원하는 경기도 청년 노동자 통장!

경기도 거주 가계통장소득 100% 이하 가구의 청년이 금융·직업교육·취업지원 등을 위해 경기도 예산 총액으로 680억원이 적립되는 통장으로 경기도 청년 노동자의 노동의 지출입금을 편리하게 기록·관리할 수 있도록 지원하여 청년의 다양한 어려움을 지원합니다.

* 총지급액은 580만원으로, 480만원 현금과 100만원 지역채권으로 지급

- 우리나라 지역 중 청년 유입 수와 청년인구 수가 가장 많은 경기도의 데이터가 2030 청년층의 목소리를 잘 반영할 수 있을 것이라 판단하여 경기도의 청년 복지 정책 중 '경기도

청년 노동자 통장'의 신청 데이터를 분석하기로 결정하였다. '청년 노동자 통장'은 경기도 거주 중위소득 이하의 청년들을 선발해 정기적으로 저축을 할 시 추가적으로 580만원의 지원금을 주는 제도이다.

● 활용한 데이터

구분	원천 데이터셋 링크	주요 컬럼
경기도 일자리 청년통장 정보	https://www.bigdata-region.kr/#/dataset/6d97268b-e314-4197-a66d-baaf0012eba1	신청년도, 저축목적명, 신청사유 내용, 연령, 성별코드, 가구원수
경기도 일자리 청년통장선발정보	https://www.bigdata-region.kr/#/dataset/9a362cbc-33a7-49aa-a0b0-eeefed39bd12	소득인정금액

- 본 분석에서는 '경기지역경제포털'에서 제공하는 청년 통장 제도 관련 데이터셋인 '일자리 청년통장정보'와 '경기도 일자리 청년통장선발정보'를 사용하였다. 이 데이터는 2017년부터 2020년까지 '경기도 청년 노동자 통장'의 신청자가 기입한 정보를 담고 있다.

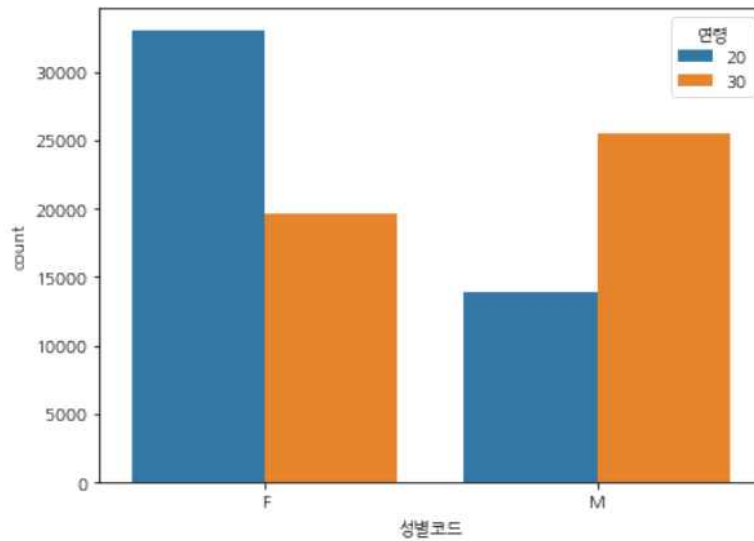
● 데이터 주요 컬럼

컬럼명	설명	비고
신청년도	청년통장 신청년도	4개 연도 (2017~2020년)
저축목적명	신청자 저축목적 분류	6개 저축목적 카테고리, 예: 창업자금, 결혼자금, 주거자금
신청사유내용	청년통장 신청사유	신청자의 정책 신청사유에 대한 텍스트 데이터
연령	신청자 연령대	신청자가 속한 연령대(10단위), 20대와 30대만 취급
성별코드	신청자 성별	남성, 여성
가구원수	신청자가 속한 가구원수	5인 이상 가구는 한 카테고리로 통합 (5인 이상)
소득인정금액	신청자 가구의 소득인정금액	5개 소득인정금액 카테고리, 예: ~180만원, 180~270만원

- 데이터 전처리 과정에서 신청인원을 구분할 수 있는 '청년통장정보번호'를 기준으로 두개의 데이터 파일을 병합하고 연 단위 신청이란 점을 고려하여 연도 컬럼을 추가하였다. 또한 신청 시간과 관련 없는 컬럼들이 모두 같은 값을 가지는 경우나 동일한 신청사유내용을 가진 경우는 중복 데이터로 판단하고 제거하였다.

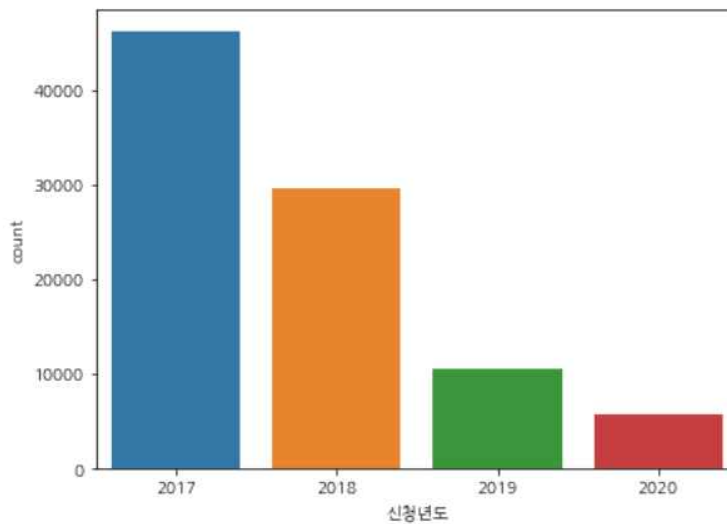
- 추가로 '소득인정금액' 컬럼의 경우, 4분위수를 기준으로 하여 '~180만원', '180~270만원', '270~401만원', '401~1000만원', '1000만원~'과 같이 5개의 구간으로 구분하였다. 또한 가구원수 컬럼도 5인 이상 가구의 데이터는 많지 않았으므로 이러한 경우를 모두 합하여 '5인 이상'이라는 하나의 카테고리로 지정하였다. 이처럼 전처리를 거친 후, 최종적으로 도출된 92,182개의 데이터로 분석을 수행하였다.

2) EDA (탐색적 데이터 분석)



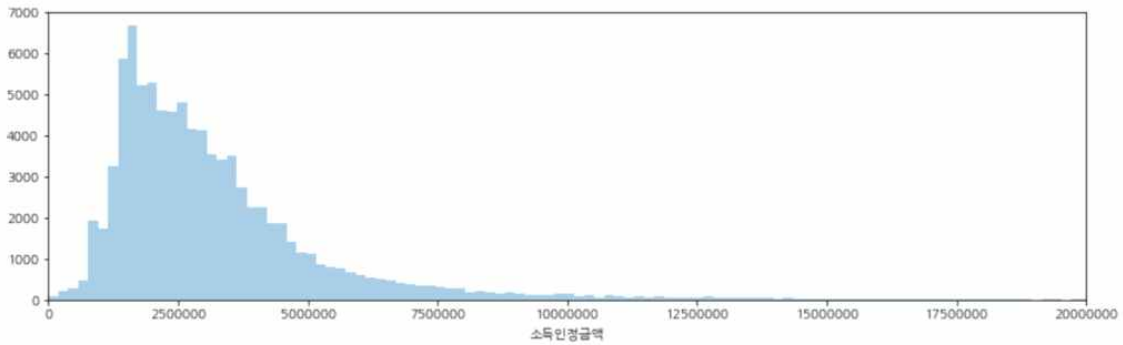
[신청인구의 연령과 성별 분포]

- 남성(M)보다는 여성(F)이, 30대보다는 20대가 더 많은 신청 건수를 기록하였다. 여성은 30대보다 20대가 많이 신청, 남성은 20대보다 30대가 더 많이 신청하였다.



[연도별 신청건수]

- 연도별 신청건수는 해가 지날수록 감소하는 양상을 보이는데, 2017년에는 4만 건 이상이던 신청건수가 2020년에는 1만 건 이하인 것을 확인할 수 있다. 2020년에 '청년 노동자 통장'의 모집 인원이 9천명이었다는 것을 감안하면 실제 신청인원이 줄어든 것이 아니라 데이터가 일부 누락된 것이라고 추정할 수 있다.



[소득인정금액 분포]

- '소득인정금액'이란 월 단위로 계산된 소득평가액에 재산의 소득환산액을 합친 금액이며 가구 단위로 계산된다. 본 데이터에서 소득인정금액의 평균은 378만 원, 중위값은 270만원을 기록하였다. 또한 하위 25%는 180만 원, 상위 25%는 401만 원인 것으로 나타났다.

<가구원수:1>
 해당 인원: 13,667명
 평균 소득인정금액: 1,644,692.58원

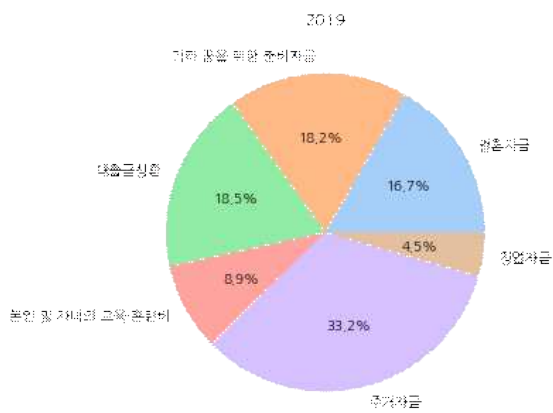
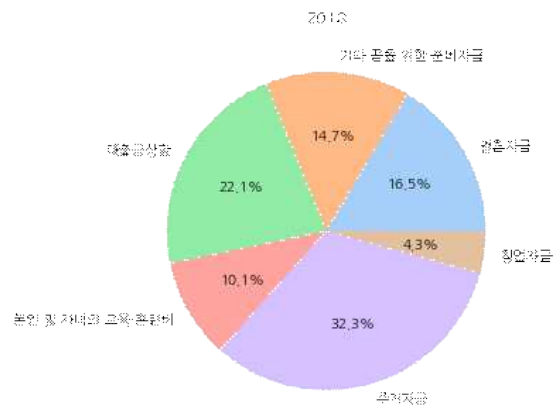
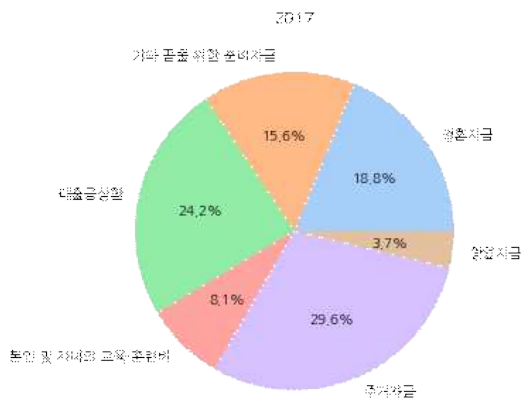
<가구원수:2>
 해당 인원: 16,123명
 평균 소득인정금액: 2,745,475.77원

<가구원수:3>
 해당 인원: 27,758명
 평균 소득인정금액: 3,783,338.05원

<가구원수:4>
 해당 인원: 26,591명
 평균 소득인정금액: 4,909,110.07원

<가구원수:5 이상>
 해당 인원: 7,880명
 평균 소득인정금액: 5,820,863.13원

- 2021년 기준 가구원 수별 중위소득과 본 데이터의 가구원 수별 소득인정금액을 비교하면 1, 2, 3인 가구에서 평균 소득인정액이 모두 기준 중위소득 100%보다 낮은 것으로 나타났다. 또한, 가구원 수가 늘어날수록 평균 102만 원의 소득인정금액이 증가하는 것을 확인할 수 있으며 가구원 수는 3인과 4인 가구가 가장 많은 것으로 드러났다. (전체의 약 46%)



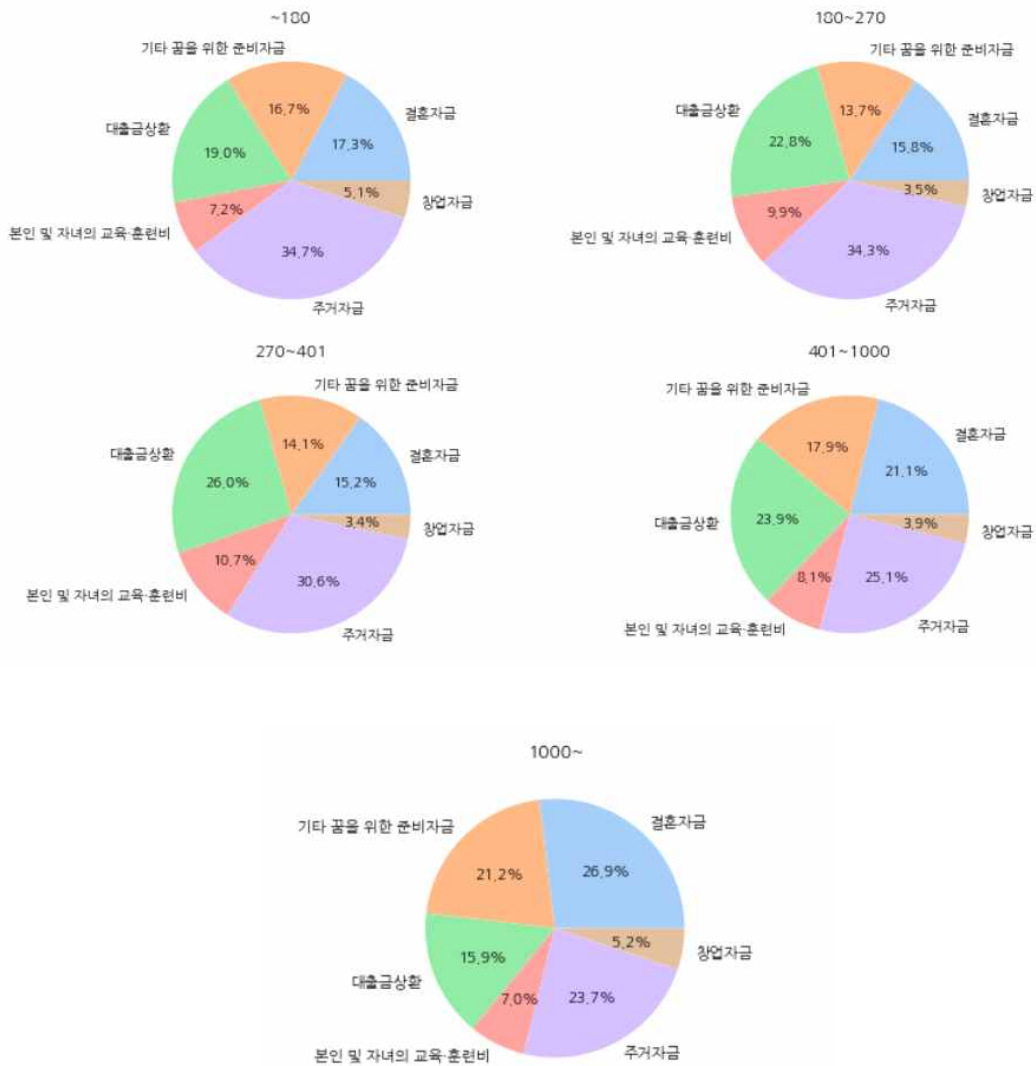
[신청년도별 저축목적 비율의 변화]

- 2017년부터 2020년까지 4년간 주거자금과 꿈을 위한 준비자금의 비중이 지속적으로 늘어난 것을 확인할 수 있다.
(주거: 29.6% → 35.6%, 꿈: 15.6% → 20.1%)
- 반면, 결혼과 대출금 상환에 대한 비중은 계속해서 줄어들었다.
(결혼: 18.8% → 12.6%, 대출금: 24.2% → 16.4%)



[성별&연령별 저축목적 비율의 차이]

- 20대 여성과 20대 남성의 저축목적의 비율은 주거자금, 기타 꿈을 위한 준비자금, 대출금상환, 결혼자금, 본인 및 자녀의 교육·훈련비, 창업자금 순으로 높다.
- 30대 여성의 저축목적의 비율은 주거자금, 대출금상환, 결혼자금, 기타 꿈을 위한 준비자금, 본인 및 자녀의 교육·훈련비, 창업자금 순으로 높다.
- 30대 남성의 저축목적의 비율은 주거자금, 대출금상환, 본인 및 자녀의 교육·훈련비, 결혼자금, 기타 꿈을 위한 준비자금, 창업자금 순으로 높다.
- 20대에서 성별에 따른 저축목적의 차이는 크지 않고, 20대와 30대는 기타 꿈을 위한 준비자금이 10퍼센트 이상 차이난다.
- 30대 여성은 저축목적에서 결혼자금이 차지하는 비율이 다른 집단보다 높다.
- 30대 남성은 저축목적에서 본인 및 자녀의 교육·훈련비가 차지하는 비율이 다른 집단보다 높다.



[소득인정금액별 저축목적 비율의 차이]

- 소득인정금액이 커질수록 저축목적에서 주거자금의 비율은 줄어들고, 결혼자금의 비율은 늘어난다.
- 이처럼 같은 청년층이라 해도 연령, 성별, 소득수준 등의 여러 인구 통계학적 변수에 따라 니즈의 차이가 존재한다. 하나의 변수를 선택해 단편적인 분석을 하는 것보다 여러 변수들을 종합하여 보는 것이 실제 청년들의 니즈 차이를 파악하는 데 중요할 것으로 판단하였다. 단, 모든 변수를 고려할 경우, 100가지의 경우의 수가 나오기 때문에 비슷한 배경의 연령, 성별, 소득수준, 가구원 수를 기반으로 클러스터링을 한 결과로 분석을 시행하였다.

[목적 : 본인 및 자녀의 교육·훈련비] : 대출금상환 or 양육비로 사용하고 싶습니다.!!전세난으로 인하여 부득히 집을 힘들게 마련하였는데 대출금상환도 힘듭니다.!!조금이나마 도움을 받고 싶어서 신청합니다.

[목적 : 주거자금] : 결혼자금 및 창업; 학자금의 밀천으로 저축하고 싶습니다. 타은행의 적금제도는 금리비율이 매우 낮아; 우연히 알게된 경기도 일하는 청년통장에 매달 10만원씩 성실히 저축하여 향후 좋은 교육자가되어 사회에 다시 환원하는 좋은일에 쓰겠습니다.

[신청사유 예시 1]

[목적 : 주거자금] : 월세 집에서 소득이 없으신 아버지를 모시고 살고 있으며; 저의 현재 급여만으로는 생활비와 월세를 충당하는데 부담이 있습니다. 2020년까지 저축액 5000만 원을 마련할 계획이며; 학자금대출 상환; 전세자금에 사용; 기타 이사비용에 충당할 계획입니다. 이번 청년통장 지원을 통해 전세집 보증금을 마련하는데 도움을 받고; 저의 목표를 실천하고 싶습니다.

[목적 : 결혼자금] : 결혼자금 혹은 주거자금으로 계획을하며 이 신청서를 넣어보려고 하고 있습니다. 추가적으로 자녀계획으로도 쓸일이 생긴다면은 교육바로 쓰려고 생각하고 있습니다.

[신청사유 예시 2]

- 이어서 위의 이미지는 데이터 중 '신청사유내용' 컬럼의 실제 예시이다. [예시1]처럼 실제 목적과 '저축목적'의 선택지가 다른 경우도 있고, [예시 2]처럼 신청자의 저축목적이 하나가 아닌 여러 개인 경우도 존재한다. 따라서 실제로 청년들이 원하는 것이 무엇인지 더욱 상세히 파악할 수 있으며 정책 신청 동기에 대한 추가적인 정보를 얻을 수 있다는 점을 고려해 '신청사유내용' 컬럼을 텍스트마이닝을 통해 알아볼 필요가 있다고 판단하였다.

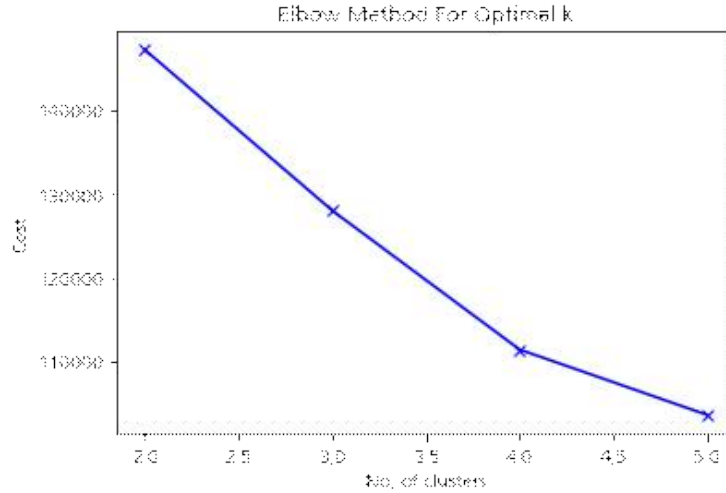
3) 클러스터링

연령	성별코드	소득인정금액수정	가구원수 수정
20	F	401~1000	3
20	F	~180	1
30	F	270~401	4
20	F	401~1000	4
30	F	~180	2

[최종 선별 변수 예시]

- 클러스터링을 위해 최종적으로 선별한 변수는 EDA를 통해 청년들의 특성 파악에 중요하다고 판단된 연령, 성별, 소득수준, 가구원 수이며, 변수 형태는 위의 예시와 같이 범주형 자료 형태이다. 따라서, 범주형 변수에 적용하는 클러스터링 방법인 '**K-Modes Clustering**'을 사용하였다. '**K-Modes Clustering**'은 범주형 개체에 대해 단순 일치 비유사성

측정을 사용하고, 최빈값으로 클러스터 수단을 대체하며 주파수 기반 방법을 사용하여 문제를 해결할 모드(modes)를 찾는 기법이다. 이어서 'Elbow Method'를 통해 최적의 클러스터 수를 찾아 최종적으로 4개의 클러스터를 도출하는 것으로 결정하였다.



* Elbow method: Cluster 간의 거리의 합을 나타내는 inertia가 급격히 떨어지는 구간이 생기는데 이 지점의 K값을 최적의 군집의 개수로 사용

구분		Cluster 0	Cluster 1	Cluster 2	Cluster 3
연령	20	33284	3413	10166	
	30	6134	27150	4109	7763
소득인정금액	~180	2179	5317	14054	861
	180~270	13286	6164		4187
	270~401	9873	12994	99	
	401~1000	11778	5141	89	2194
	1000~	2302	957	33	521
가구원수	1	1168	3138	8825	536
	2	4317	6048	3339	2419
	3	8102	14341	1565	3750
	4	21339	5252		
	5 이상	4492	1784	546	1058
클러스터 특징		20대 여성& 남성 중위소득 가족과 함께 거주	30대 남성 중위소득 가족과 함께 거주	20대 여성 저소득 1인가구	30대 여성 중위소득 가족과 함께 거주

- 앞서 설명한 'K-modes Clustering'으로 도출한 각 클러스터의 데이터는 위와 같다. 각 클러스터별로 높은 수치를 기록한 데이터를 바탕으로 하여 가장 눈에 띄는 특징들로 클러스

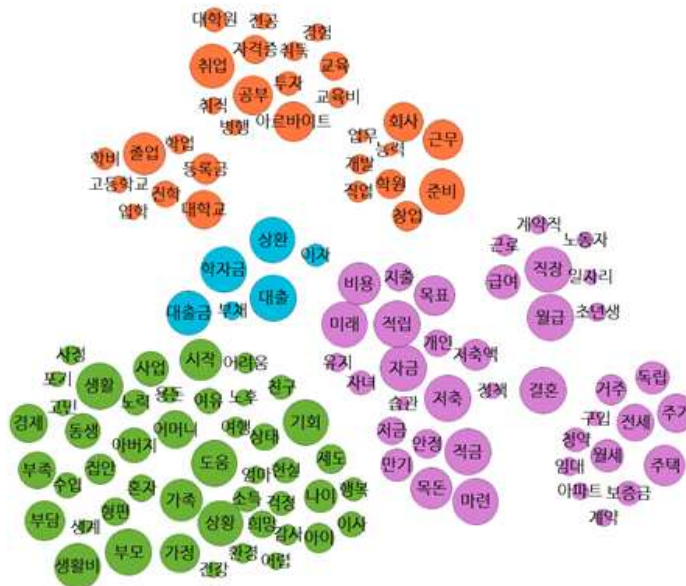
터를 구분하였다. 그 결과, 클러스터 0은 '20대 여성 & 남성, 중위소득, 가족과 함께 거주'라는 특징을 갖고 있으며 클러스터 1은 '30대 남성, 중위소득, 가족과 함께 거주'라는 특징을 보인다. 이어서 클러스터 2는 '20대 여성, 저소득, 1인 가구'의 특징을 갖고 있으며 클러스터 3의 경우, '30대 여성, 중위소득, 가족과 함께 거주'라는 특징을 확인할 수 있다.

- 이처럼 도출된 4개의 클러스터별 니즈와 특성을 좀 더 세밀하게 파악해보고자 단어 네트워크 분석과 DMR 토픽 모델링 기법을 활용하여 텍스트마이닝을 실행하였다.

4) 텍스트마이닝

① 클러스터별 단어 네트워크 분석을 통한 청년 페르소나 제시

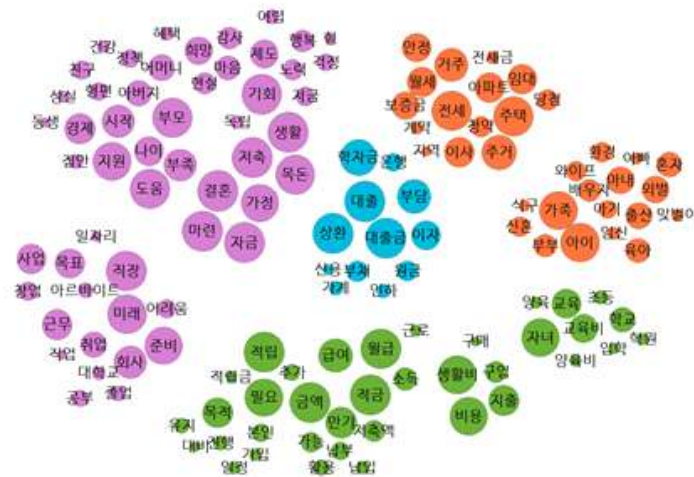
- Textrank를 기반으로 클러스터별 '신청사유' 컬럼에 대한 주요 키워드를 추출한 뒤, 동시 출현한 단어들을 활용하여 단어 네트워크 분석을 진행하였다. 각 클러스터별로 도출된 단어 네트워크 맵을 바탕으로 정책 수요자인 청년 그룹별 페르소나를 제시해 보고자 한다. 여기서 '페르소나'란 '어떤 제품 혹은 서비스를 사용할 만한 목표 인구 집단 안에 있는 다양한 사용자 유형들을 대표하는 가상의 인물'을 뜻한다.



[클러스터0 단어 네트워크 맵]

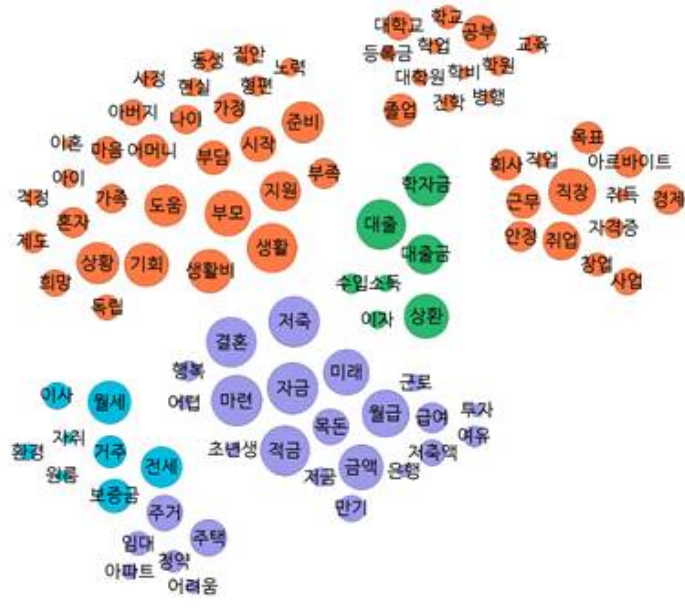
- 먼저 클러스터 0의 경우, 공부, 대학교, 졸업 등 학업과 관련된 단어가 등장하는 것을 볼 수 있으며 취업, 직장, 월급, 초년생 등 사회 초년생으로서 직장생활을 시작하는 모습을 엿볼 수 있는 단어들이 나타난다.

- 또한 독립, 주거, 전세, 월세, 보증금 등의 단어가 등장하여 독립을 위한 주거자금 마련에 관심이 있다고 추측할 수 있다. 그 외로는 학자금 대출 상황, 가족 생활비, 미래를 위한 저축 등에 대한 단어들이 함께 등장하고 있다.
- 이러한 단어 네트워크 분포를 고려하면 '20대 여성 & 남성, 중위소득, 가족과 함께 거주'라는 특징을 가지고 있는 클러스터 0은 '독립을 꿈꾸는 사회초년생'이라는 페르소나를 부여할 수 있다.



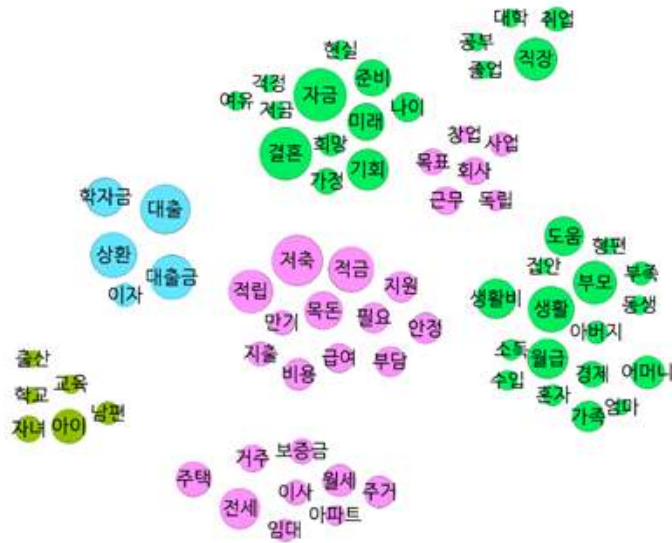
[클러스터1 단어 네트워크 맵]

- 이어서 클러스터1의 단어 네트워크맵을 살펴보면 앞선 클러스터와 동일하게 학자금 및 대출금 상환, 주거 등에 대한 단어도 빈번하게 등장하지만 가정, 아내, 외벌이, 아이, 생활비 등 가족과 관련한 단어들이 많이 등장하는 것이 특징적이다. 이때, 자녀, 교육비, 양육비, 입학 등 자녀 교육과 관련된 단어들 또한 함께 나타나는 것을 볼 수 있다.
- 그 외로 직장, 사업, 회사, 근무, 목표, 미래 등 사회초년생 시기를 지나 좀 더 미래에 대한 목표를 가지고 사회생활을 이어나가고 있는 모습을 엿볼 수 있는 단어들을 확인할 수 있다.
- 단어 네트워크맵을 통해 도출된 단어들과 '30대 남성, 중위소득, 가족과 함께 거주'라는 특징을 종합해 보았을 때 클러스터1은 '미래를 준비하는 초보 아빠'라고 이름 붙일 수 있다.



[클러스터2 단어 네트워크 맵]

- '20대 여성, 저소득, 1인 가구'라는 특성을 가진 클러스터2는 공부, 대학, 취업, 창업 등 진로와 관련한 단어들이 등장한다는 면에서 앞선 클러스터0과 비슷한 특징을 보이나 몇가지 측면에서 차이를 보인다.
- 먼저 주거와 관련하여 자취, 원룸, 이사, 환경, 어려움 등의 단어가 등장하는 것을 미루어 보아 클러스터2에 해당하는 청년들은 이미 독립을 이루기는 했으나 주거 환경을 좀 더 개선하고 싶은 니즈를 엿볼 수 있다.
- 또한 가족과 떨어져 사는 1인 가구인 만큼 가족의 사연과 관련된 단어의 등장이 클러스터0보다 적은 것을 확인할 수 있다.
- 따라서 클러스터2는 '홀로서기 연습 중인 20대 여성'이라고 해석할 수 있다.



[클러스터3 단어 네트워크 맵]

- 마지막으로 '30대 여성, 중위소득, 가족과 함께 거주'라는 특징을 가진 클러스터3의 경우, 주거, 직장, 대출 상환 등 모든 클러스터에서 동일하게 등장하는 단어를 제외하면 자녀, 아이, 출산, 교육 등 자녀 양육과 관련된 단어 군집이 가장 눈에 띈다.
- 또한 결혼, 가정, 미래, 준비 등의 단어를 고려하였을 때, 결혼 후 가정을 꾸리고자 하는 희망이 엿보인다.
- 위의 특징들을 고려하였을 때, 클러스터3은 '행복한 가정을 꿈꾸는 예비 신부/엄마'의 페르소나라고 할 수 있다.

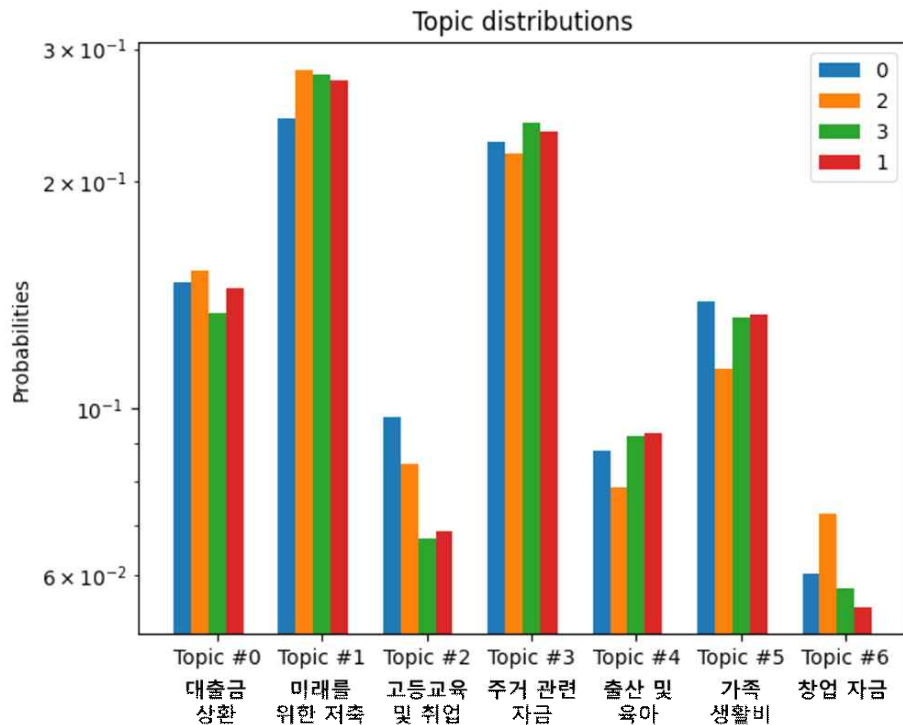
② 클러스터별 DMR 토픽 모델링

- DMR 토픽 모델링은 LDA(잠재 디리클레 할당, Latent Dirichlet allocation)를 기반으로 하여 메타데이터별 주제 분포를 추정하는 기법이다. 앞서 도출한 클러스터를 메타데이터로 하여 총 7개 토픽의 특징적인 키워드들을 추출하고, 각 토픽이 클러스터를 구성하는 비율을 구했다.

토픽0 대출금 상환	토픽1 미래를 위한 저축	토픽2 고등교육 및 취업	토픽3 주거 관련 자금	토픽4 출산 및 육아	토픽5 가족 생활비	토픽6 창업 자금
대출	저축	공부	저축	아이	부모	창업
상환	미래	대학	적금	자녀	어머니	꿈
학자금	꿈	졸업	전세	교육비	아버지	사업
대출금	부모	꿈	적립	생활	가족	운영
저축	적금	취업	주택	가족	동생	카페
이자	도움	아르바이트	월세	전세	생활	가게
적립	지원	대학교	주거	아내	생활비	학원
생활비	생활	대학원	보증금	교육	도움	목표
적금	이번	진학	월급	출산	집안	회사
빚	직장	자격증	거주	육아	엄마	기술

[토픽별 주요 키워드]

- 위의 표와 같이 토픽별로 추출된 주요 키워드들을 바탕으로 각 토픽이 어떤 주제를 갖고 있는지 유추하였으며, 그 결과 '대출금 상환', '미래를 위한 저축', '고등교육 및 취업', '주거 관련 자금', '출산 및 육아', '가족 생활비', '창업자금'이라는 토픽을 도출할 수 있었다.



[토픽별 클러스터 비율]

- 위의 토픽별 클러스터 비율 그래프를 살펴보면 모든 클러스터에서 공통적으로 '미래를 위한 저축'과 '주거 관련 자금' 토픽이 높은 언급도를 보인다는 것을 알 수 있다. 그 다음으로는 '대출금 상환', '가족 생활비', '출산 및 육아'가 뒤따랐다.
- 토픽0, 1, 3, 4, 5는 클러스터별로 크게 주요 키워드의 언급의 차이가 없었으나, 토픽2(고등교육 및 취업)과, 토픽6(창업 자금)의 경우 클러스터 별 편차가 컸다. '고등교육 및 취업' 토픽의 경우, 30대 청년이 많은 클러스터1과 클러스터3에 비해 구성인원이 20대인 클러스터 0('독립을 꿈꾸는 사회초년생')과 클러스터 2('홀로서기 연습 중인 20대 여성')에서 눈에 띄게 언급도가 높았다. 한편, 창업과 관련된 토픽6은 가족 구성원이 많은 클러스터 0, 1, 3보다 클러스터 2에서 관련 단어의 언급도가 높았다.

클러스터	Cluster 0	Cluster 1	Cluster 2	Cluster 3
페르소나	독립을 꿈꾸는 사회 초년생	미래를 준비하는 초보 아빠	홀로서기 연습 중인 20대 여성	행복한 가정을 꿈꾸는 예비 신부/엄마
높은 니즈	고등교육 및 취업 가족 생활비	미래를 위한 저축 주거 관련 자금 출산 및 육아 가족 생활비	대출금 상환 미래를 위한 저축 고등교육 및 취업 창업	미래를 위한 저축 주거 관련 자금 출산 및 육아
낮은 니즈	미래를 위한 저축	고등교육 및 취업 창업 자금 대출금 상환	주거 관련 자금 출산 및 육아 가족 생활비	고등교육 및 취업 창업 자금 대출금 상환

[클러스터별 니즈 비교]

- 클러스터별로 각 토픽에 대한 니즈를 좀 더 자세히 살펴보기 위해 상대적인 비교를 실시하였다. 각 토픽별로 비율 1, 2위를 차지한 클러스터는 해당 토픽에 높은 니즈를 갖고 있다고 판단하였으며, 반대로 3, 4위를 차지한 클러스터의 경우에는 낮은 니즈를 갖고 있다고 판단하였다.
- 그 결과, 위의 표와 같은 결과가 도출되었으며 대부분 클러스터별 페르소나와 부합하는 니즈가 도출되었다. 다만 '홀로서기 연습 중인 20대 여성' 페르소나를 가진 클러스터2에서 창업에 대한 니즈가 다른 클러스터들에 비해 월등히 높게 도출된 점이 주목할 만한 점이라고 생각되었다. 특히 전체 DMR 그래프를 살펴보면 각 토픽에 대한 클러스터별 비율이 대부분 고르게 분포하고 있기 때문에 창업 토픽에서 클러스터2만 높게 도출된 결과가 더욱 두드러진다.
- 따라서 이러한 특성을 고려하여 결론에서 '홀로서기 연습 중인 20대 여성' 페르소나를 가진 클러스터2에 맞는 **창업 정책 제안**을 다루어 보고자 한다.

5) 결론 및 의의

① 청년 페르소나를 기반으로 한 정책 타겟층 도출



독립을 꿈꾸는
사회 초년생



홀로서기 연습 중인
20대 여성



미래를 준비하는
초보 아빠



행복한 가정을 꿈꾸는
예비 신부/엄마

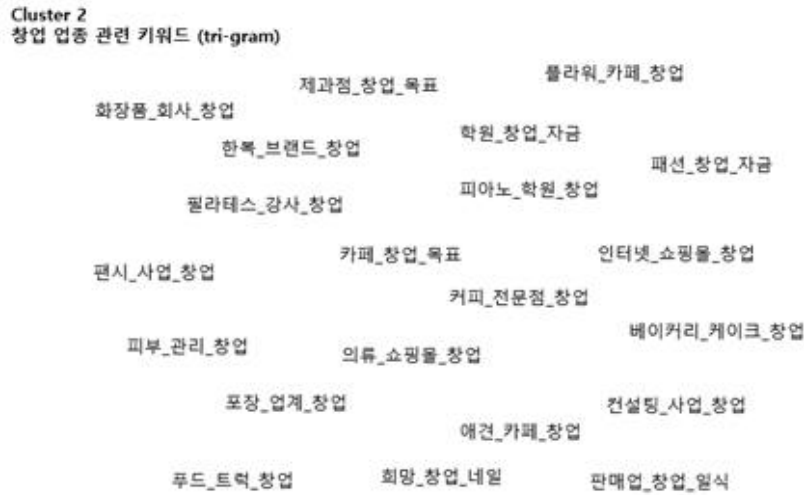
- 앞서 '경기도 청년 노동자 통장'에 지원한 차상위계층 청년들의 인구통계학적 특징을 고려하여 4개의 클러스터를 도출하였고, 단어 네트워크 맵을 활용해 각 클러스터별 청년 유형을 대표하는 페르소나를 부여하였다. 이와 같이 데이터 분석을 기반으로 도출된 페르소나는 정책 수요자 청년들이 어떤 배경과 니즈를 가진 인물인지 명확하게 파악하는 데 도움을 준다는 점에서 의미를 가진다. 또한 정책 기획이나 홍보 등에 활용할 수 있다는 장점이 있다.
- 먼저 청년 대상 정책을 기획할 때, 청년 페르소나를 참고하여 수요자 맞춤형 정책을 기획할 수 있다. 예를 들어 '미래를 준비하는 초보 아빠' 페르소나에 속하는 청년의 경우, 가정, 육아에 대한 책임감과 미래에 대한 고민이 많은 30대 기혼 남성의 특징을 가지고 있으므로 가족 생활비 관련 정책, 자녀와 함께하는 프로그램, 자기계발 등에 대한 지원을 맞춤형으로 기획할 수 있다.



[경기도 복지 정책 홍보 페이지 모습]

- 그뿐만 아니라 정책을 소개하거나 홍보할 때, 도출된 페르소나를 활용하여 청년들의 관심을 유도할 수 있다. 위의 이미지는 현재 경기도의 복지 정책을 홍보하는 웹사이트의 실제 모습인데 이와 비슷하게 페르소나를 활용한 청년 정책 홍보가 가능하다. '독립을 꿈꾸는 사회 초년생', '미래를 준비하는 예비 아빠', '홀로서기 연습 중인 20대 여성', '행복한 가정을 꿈꾸는 예비 신부/엄마'와 같이 페르소나별로 필요한 정책을 소개한다면 청년들이 자신에게 맞는 정책이 무엇인지 찾는 과정이 더욱 편리해 질 수 있다. 예를 들어 결혼과 육아, 내집마련에 관심이 많은 30대 여성이라면 '행복한 가정을 꿈꾸는 예비 신부/엄마' 카테고리에 속하는 정책들을 클릭해 볼 것이다.

② 20대 여성 창업에 대한 정책 현황 분석 및 개선점 제안



[클러스터2 창업 업종 관련 tri-gram 키워드]

- 앞서 DMR 그래프를 통하여 '홀로서기 연습 중인 20대 여성' 페르소나를 가진 클러스터2에서 창업에 대한 니즈가 다른 클러스터들에 비해 두드러지는 것을 확인할 수 있었다. 따라서 이에 대해 좀 더 자세히 알아보기 위하여 클러스터2의 신청사유 컬럼에서 연속된 3개의 키워드를 추출하는 tri-gram 분석을 진행하였다. 그 결과, 실제로 제과점, 학원, 애견카페, 커피 전문점, 쇼핑몰, 네일샵 등 다양한 업종에 대한 수요가 있는 것이 확인할 수 있었다.
- 그러나 창업과 관련된 청년 정책들을 살펴보면 이러한 현실 니즈와 맞지 않는 내용으로 진행되는 것들이 많다. 그에 대한 예시로 경기도 군포시에서 진행되었던 청년창업 지원 교육의 경우, 주로 스타트업과 비즈니스 모델 개발, IR 피칭 등 스타트업 창업과 관련된 내용이 주를 이루고 있다. 이는 앞서 tri-gram 분석에서 살펴본 것과 같이 자영업 창업 수요가 많은 20대 여성 창업자들의 현실적인 니즈와 거리가 있는 것을 확인할 수 있다.
- 혹은 창업에 대한 추가적인 교육 없이 재정적 지원을 주로 하는 청년창업 지원 정책이 주를 이룬다. 그러나 최근 윤석현 금융감독원장이 한 행사에 참가하여 "청년층의 자영업 진출이 증가하고 있지만 폐업률이 높아 적극적인 지원이 필요하다"며 지적한 바와 같이 청년 창업에 대한 니즈가 증가하고 있으나 전 연령대에서 가장 높은 단기폐업 비율을 보이고 있다. 따라서 진정한 의미의 청년 창업 지원을 위해서는 단순한 재정 지원뿐만 아니라 창업에 대한 기본적인 교육도 함께 진행되어 폐업률을 낮추고 장기적으로 안정적인 운영을 할 수 있도록 지원해야 한다.

모범 사례

경기도 하남시 '일마련 청년창업 지원 사업'

□ 지원내용

- 창업역량교육 - 기본교육 및 전문교육(5일) = 26시간 이상
- 창업지원 - 창업교육, 컨설팅, 평가 등을 통해 지원대상자로 선정

□ 세부지원내역

- 인테리어 : 최대 1,000만원
- 임대료 : 최대 월 70만원(아무상 30%)
 - 인테리어 및 임차료는 창업비용에 따라 평가·평가를 통해 결정
 - 선정세월로 지원 금액이 달라질 수 있음

※, 지원가능액 및 신청자격에 따라 다르므로 지원 가능여부 확인 후 지원신청을 하시기 바랍니다.
 ※, 지원가능액 및 신청자격에 따라 다르므로 지원 가능여부 확인 후 지원신청을 하시기 바랍니다.
 ※, 지원가능액 및 신청자격에 따라 다르므로 지원 가능여부 확인 후 지원신청을 하시기 바랍니다.

하남 일마련 청년창업자, '여름다운 기부'로 보답

2023. 07. 26

하남 일마련 청년창업자 10명, 여름달기 기부금 100만원



- 앞서 지적한 실제 수요와 동떨어진 창업 교육 내용, 재정적 지원에 치중된 창업 지원 프로그램의 현실을 고려하였을 때, 경기도 하남시의 '일마련 청년창업 지원 사업'은 부족한 부분들을 보완한 모범사례라고 할 수 있다. 창업에 대한 기본 교육을 진행하며 창업 비용에 대한 지원도 함께 이루어지고 있기 때문이다. 실제로 이 정책을 통해 케이크샵 창업을 이루어낸 여성 청년 창업가가 저소득 아동에게 기부를 하는 선순환이 이루어진 모습까지 확인할 수 있다. 따라서 이러한 모범사례에 앞선 데이터 분석 결과를 더하여 20대 여성 창업가를 주요 타겟으로 한 청년창업 지원 정책의 방향성을 제안하고자 한다.



- 본 제안이 기존 정책과 차별화된 점은 토픽모델링(DMR)을 통해 20대 저소득층 여성의 창업 니즈가 다른 청년 그룹과 비교했을 때 두드러진다는 점을 파악하여 청년 창업 정책의 타겟을 명확히 제시한 것이다.
- 또한 창업 지원 프로그램의 내용에 사업 운영에 대한 교육뿐만 아니라 창업 아이템에 대한 교육을 추가했다는 점에서 의미가 있다. 이를 통하여 창업에 대한 욕구가 있으나 아직 실질적인 스킬이 부족한 청년들이 도움을 받을 수 있다. 특히 앞선 tri-gram 분석을 통하여 '홀로서기 연습 중인 20대 여성' 페르소나를 가진 클러스터2에서 수요가 높은 창업 아이템들을 도출하여 이를 반영하였다. 따라서 향후 청년을 대상으로 한 창업 지원 정책을 기획할 시, 본 분석 내용을 참고하여 20대 여성 청년들의 창업 활동을 독려할 수 있다.
- 결론적으로 본 분석의 전반적인 의의는 청년 정책에 지원한 2030 세대들이 실제로 어떤 특성 혹은 배경을 갖고 있는지, 어떤 부분에 대한 지원을 필요로 하는지 데이터를 통해 파악했다는 것에 있다. 특히 클러스터링과 텍스트마이닝 기법을 활용하여 청년들의 페르소나를 도출하고, 다른 클러스터에 비해 두드러지는 20대 여성 청년들의 창업 니즈를 발견하여 향후 정책 기획 및 홍보 등에 활용이 가능하다는 점에서 효과적이라고 할 수 있다.